¿Qué es la escala de Likert y cómo utilizarla?

Antecedentes de la escala de Likert

La **escala de Likert** es una de las herramientas más utilizadas por los investigadores de mercado cuando desean evaluar las opiniones y actitudes de una persona. Existen varios tipos de escalas de medición enfocadas directamente a medir las actitudes de las personas, entre ellas, una de las más utilizadas es de la que hablaremos a continuación.

Qué es la escala de Likert

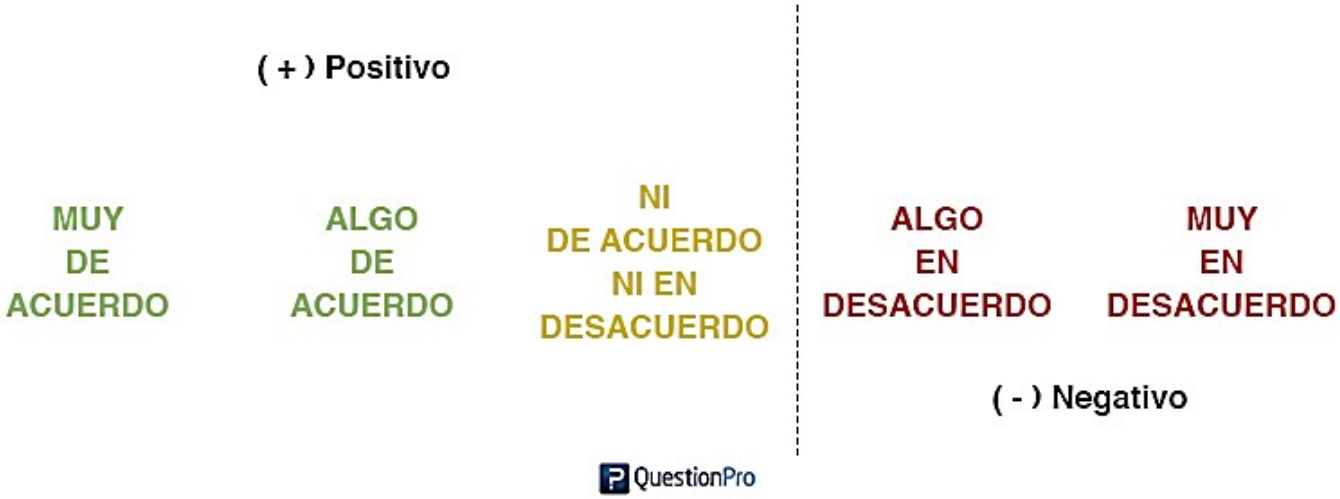
Empecemos con el nombre de la escala, el cual tiene su origen debido al psicólogo Rensis Likert. Likert distinguió entre una escala apropiada, la cual emerge de las respuestas colectivas a un grupo de ítems (pueden ser 8 o más), y el formato en el cual las respuestas son puntuadas en un rango de valores.

Técnicamente, una escala de Likert hace referencia al último. La diferencia de estos dos conceptos tiene que ver con la distinción que Likert hizo entre el fenómeno que está siendo investigado y las variables de los medios de captura.

La escala de Likert es uno de los tipos de escalas de medición. Es una escala psicométrica utilizada principalmente en la investigación de mercados para la comprensión de las opiniones y actitudes de un consumidor hacia una marca, producto o mercado meta. Nos sirve principalmente para realizar mediciones y conocer sobre el grado de conformidad de una persona o encuestado hacia determinada oración afirmativa o negativa.

Cuando se responde a un ítem de la escala de Likert, el usuario responde específicamente en base a su nivel de acuerdo o desacuerdo. Las escalas de frecuencia con la de Likert utilizan formato de respuestas fijos que son utilizados para medir actitudes y opiniones. Estas escalas permiten determinar el nivel de acuerdo o desacuerdo de los encuestados.

La escala de Likert asume que la fuerza e intensidad de la experiencia es lineal, por lo tanto va desde un totalmente de acuerdo a un totalmente desacuerdo, asumiendo que las actitudes pueden ser medidas. Las respuestas pueden ser ofrecidas en diferentes niveles de medición, permitiendo escalas de 5, 7 y 9 elementos configurados previamente. Siempre se debe tener un elemento neutral para aquellos usuarios que ni de acuerdo ni en desacuerdo.



Ejemplos de preguntas

Nivel de acuerdo o desacuerdo de una oración

Creo que la innovación es el elemento más importante a la hora de iniciar una propuesta.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

El sabor en la comida que consumo es lo más importante para elegir un lugar en el que comer.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Mi aportación a la empresa es siempre valiosa.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Frecuencia de realización de una actividad

Salgo de vacaciones al menos 2 veces al año.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Consumo productos para cuidado de la piel todos los días.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Importancia de un factor

La experiencia del usuario dentro de mi estrategia comercial es...

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Valoración de marca, empresa, producto o servicio

La marca de A se preocupa por el medio ambiente y la naturaleza

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

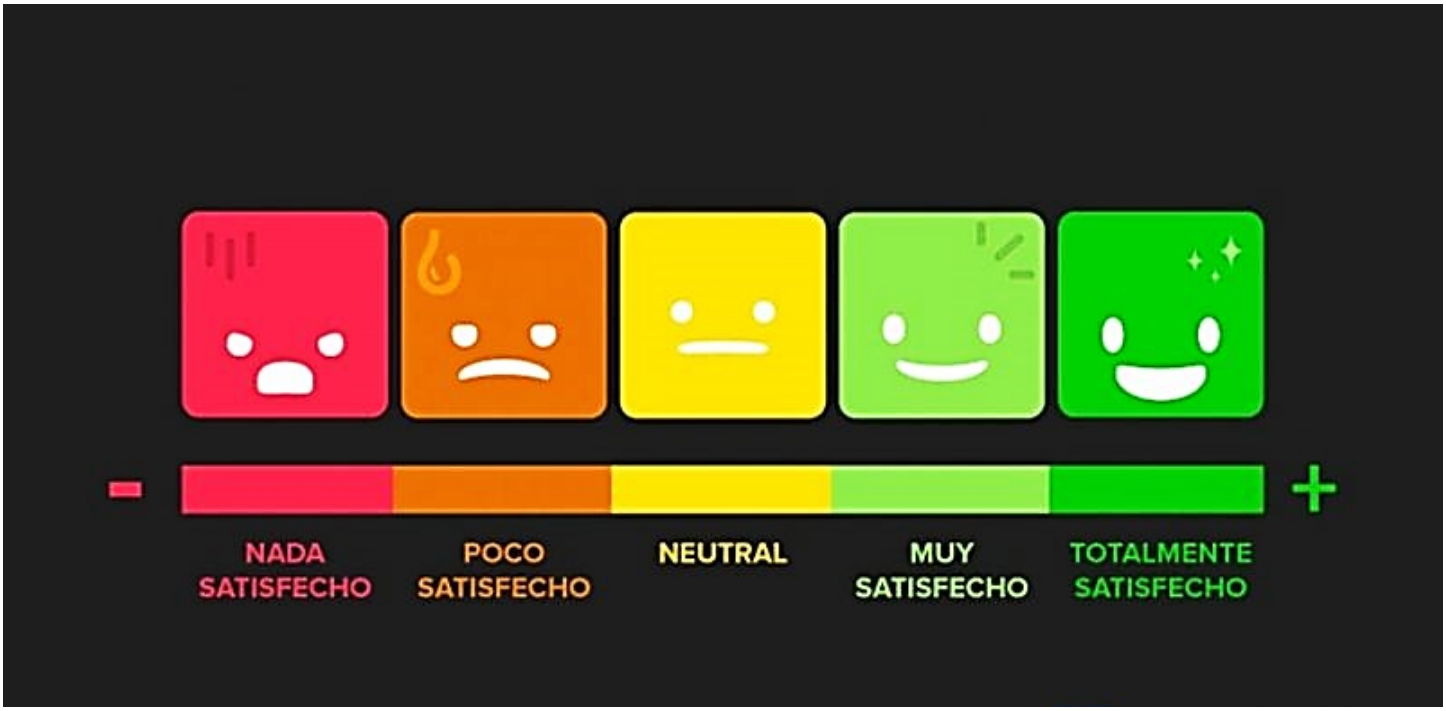
Probabilidad de realizar una acción en el futuro

Tengo planeado adquirir un automóvil en los próximos 6 meses.

Ejemplo de la escala de Likert

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo, ni en desacuerdo	De acuerdo	Totalmente de acuerdo
○	○	○	○	○

Ejemplos de escalas de Likert.


-
+
NADA SATISFECHO
POCO SATISFECHO
NEUTRAL
MUY SATISFECHO
TOTALMENTE SATISFECHO

Acuerdo	Frecuencia
· Totalmente de acuerdo · De acuerdo · Indeciso · En desacuerdo · Totalmente en desacuerdo	· Muy frecuentemente · Frecuentemente · Ocasionalmente · Raramente · Nunca
Importancia	**Probabilidad**
· Muy importante · Importante · Moderadamente importante · De poca importancia · SIn importancia	· Casi siempre verdad · Usualmente verdad · Ocasionalmente verdad · Usualmente no verdad · Casi nunca verdad

Ventajas

- Es una escala de fácil aplicación y diseño.
- Puede utilizar ítems que no tienen relación con la expresión.
- Ofrece una graduación de la opinión de las personas encuestadas.
- Muy sencilla de contestar.

Desventajas

- Existen estudios científicos que indican que existe un sesgo en la escala, ya que las respuestas positivas siempre superan a las negativas.

- También hay estudios que indican que los encuestados tienden a contestar “de acuerdo” ya que implica un menor esfuerzo mental a la hora de contestar la encuesta.
- Dificultad para establecer con precisión la cantidad de respuestas positivas y negativas.
- Si te ha quedado claro todo, es momento de crear tu primera encuesta online utilizando la escala de Likert.

Criterios fundamentales para configurar los ítems de una escala de Likert

Los ítems se deben relacionar fácilmente con las respuestas de la oración, sin importar que la relación entre ítem y oración sea evidente.

Los ítems deben de tener siempre dos posturas extremas así como un ítem intermedio que sirva de graduación entre los extremos. Es importante mencionar que a pesar de que la escala de Likert más común es la de 5 ítems, el uso de más ítems ayuda a generar mayor precisión en los resultados.

Los ítems de la escala deben ser siempre seguros y fiables. Para lograr la fiabilidad en ocasiones es necesario sacrificar la precisión de la escala.

Diferencias entre escala de Likert e ítem Likert.

Identificar cuando hablamos de una escala de Likert y un elemento tipo Likert es muy fácil. La escala de Likert está compuesta por varios ítems Likert como se muestra en la siguiente imagen.

Escala de Likert y análisis de datos

Las encuestas se utilizan constantemente para medir la calidad. Por ejemplo, las encuestas pueden utilizarse para medir la percepción del cliente sobre la calidad del producto o el desempeño de la calidad en la prestación de servicios.

Las escalas Likert son un formato de clasificación común para las encuestas. Los encuestados clasifican la calidad de alto a bajo o mejor a peor utilizando cinco o siete niveles.

Los estadísticos generalmente han agrupado los datos recogidos de estas encuestas en una jerarquía de cuatro niveles de medición:

- Datos nominales: El nivel de medición más débil que representa categorías sin representación numérica.
- Datos ordinales: Datos en los que es posible ordenar o clasificar las respuestas, pero no es posible medir la distancia.
- Datos de intervalo: En general, datos enteros en los que se pueden realizar mediciones de pedidos y distancias.
- Datos de relación: datos en los que es posible el ordenamiento significativo, distancia, decimales y fracciones entre variables.

Los análisis de datos que utilizan datos nominales, de intervalo y de relación son generalmente sencillos y transparentes. Los análisis de datos ordinales, particularmente en lo que se refiere a Likert u otras escalas en las encuestas, no lo son.

Una razón subyacente para analizar los datos ordinales como datos de intervalos podría ser la afirmación de que las pruebas estadísticas paramétricas (basadas en el teorema del límite central) son más poderosas que las alternativas no paramétricas.

Sin embargo, el tratamiento de los datos ordinales como datos de intervalo (o incluso de relación) sin examinar los valores del conjunto de datos y los objetivos del análisis puede inducir a error y tergiversar los resultados de una encuesta.

Fundamentos de las escalas de Likert

Las escalas de Likert se desarrollaron en 1932 como la familiar respuesta bipolar de cinco puntos con la que la mayoría de la gente está familiarizada hoy en día.3 Estas escalas van desde un grupo de categorías-menos a la mayoría-pidiendo a la gente que indique cuánto está de acuerdo o en desacuerdo, aprueba o desaprueba, o creer que es cierto o falso.

Realmente no hay manera incorrecta de construir una escala de Likert. La consideración más importante es incluir al menos cinco categorías de respuesta. Algunos ejemplos de grupos de categorías aparecen en la Tabla 1.

Los extremos de la escala a menudo se aumentan para crear una escala de siete puntos añadiendo “muy” a la parte superior e inferior de las escalas de cinco puntos. Se ha demostrado que la escala de siete puntos alcanza los límites superiores de la confiabilidad de la escala.4

Como regla general, Likert y otros recomiendan que sea mejor utilizar una escala lo más amplia posible. Siempre puede colapsar las respuestas en categorías condensadas, si es apropiado, para el análisis.

Con esto en mente, las escalas a veces se truncan a un número par de categorías (normalmente cuatro) para eliminar la opción “neutral” en una escala de encuesta de “elección forzada”.

Análisis, generalización a índices continuos

Como regla general, la media y la desviación estándar son parámetros no válidos para la estadística descriptiva cuando los datos están en escalas ordinales, al igual que cualquier análisis paramétrico basado en la distribución normal.

Los procedimientos no paramétricos basados en el rango, la mediana o el rango son apropiados para analizar estos datos, al igual que los métodos libres de distribución tales como tabulaciones, frecuencias, tablas de contingencia y estadísticas de chi-cuadrado.

Los modelos de Kruskall-Wallis pueden proporcionar el mismo tipo de resultados que un análisis de varianza, pero en base a los rangos y no a los medios de las respuestas. Dado que estas escalas son representativas de una medida subyacente continúa, una recomendación es analizarlas como datos de intervalos como un piloto antes de recoger la medida continua.

La Tabla 2 incluye un ejemplo de conclusiones engañosas, mostrando los resultados de la encuesta anual de Alfred P. Sloan Foundation sobre la calidad y extensión del aprendizaje en línea en los Estados Unidos. Los encuestados utilizaron una escala de Likert para evaluar la calidad del aprendizaje en línea en comparación con el aprendizaje presencial.

Mientras que el 60% de los encuestados percibe el aprendizaje en línea como igual o mejor que cara a cara, hay una minoría persistente que percibe el aprendizaje en línea como al menos algo inferior.

Si estos datos fueron analizados utilizando medios, con una escala de 1 a 5 de inferior a superior, esta separación se perdería, dando medios de 2,7, 2,6 y 2,7 para estos tres años, respectivamente. Esto indicaría un acuerdo ligeramente inferior al promedio en lugar de la distribución real de las respuestas.

Mejores prácticas para analizar los resultados de la escala de Likert

Las escalas Likert de cinco puntos se asocian comúnmente con las encuestas y se usan en una amplia variedad de entornos. Te has topado con la escala de Likert si alguna vez te han preguntado si estás totalmente de acuerdo, de acuerdo, ni de acuerdo o en desacuerdo, en desacuerdo o muy en desacuerdo sobre algo. Arriba podrás encontrar algunos ejemplos de este tipo de preguntas.

Debido a que los datos de los elementos de Likert son discretos, ordinales y tienen un alcance limitado, ha habido una larga disputa sobre la forma más válida de analizar los datos de Likert. La elección básica es entre una prueba paramétrica y una prueba no paramétrica. Los pros y los contras para cada tipo de prueba se describen generalmente como los siguientes:

- Las pruebas paramétricas, como la prueba t de 2 muestras, asumen una distribución normal y continua. Sin embargo, con un tamaño de muestra suficiente, las pruebas t son robustas a las salidas de la normalidad.
- Las pruebas no paramétricas, como la prueba de Mann-Whitney, no asumen una distribución normal o continua. Sin embargo, hay preocupaciones acerca de una menor capacidad de detectar una diferencia cuando uno realmente existe.

¿Cuál es la mejor opción? Esta es una decisión real que tienes que hacerte cuando vas a hacer un análisis de los datos de tu encuesta cuando utilizaste Likert.

A través de los años, una serie de estudios que han tratado de responder a esta pregunta. Sin embargo, han tendido a mirar un número limitado de distribuciones potenciales para los datos Likert, lo que hace que la generalización de los resultados a sufrir.

Gracias a los aumentos en la potencia de la computación, los estudios de simulación ahora pueden evaluar a fondo una amplia gama de distribuciones.

Los investigadores identificaron un conjunto diverso de 14 distribuciones que son representativas de los datos reales de Likert. El programa informático dibujó pares independientes de muestras para probar todas las combinaciones posibles de las 14 distribuciones.

En total, se generaron 10.000 muestras aleatorias para cada una de las 98 combinaciones de distribución. Los pares de muestras se analizan utilizando tanto la prueba t de dos muestras y la prueba de Mann-Whitney para comparar la eficacia de cada prueba. El estudio también evaluó diferentes tamaños de muestra.

Los resultados muestran que para todos los pares de distribuciones las tasas de error Tipo I (falso positivo) son muy cercanas a las cantidades objetivo. En otras palabras, si usted utiliza cualquiera de los análisis y sus resultados son estadísticamente significativos, no es necesario estar demasiado preocupado por un falso positivo.

Los resultados también muestran que para la mayoría de los pares de distribuciones, la diferencia entre el poder estadístico de las dos pruebas es trivial. En otras palabras, si realmente existe una diferencia a nivel de población, cualquiera de los análisis es igualmente probable que lo detecte.

Las preocupaciones sobre la prueba de Mann-Whitney que tienen menos poder en este contexto parecen ser infundadas.

Tengo una advertencia. Hay algunos pares de distribuciones específicas donde hay una diferencia de poder entre las dos pruebas.

Si realiza ambas pruebas en los mismos datos y no están de acuerdo (uno es significativo y el otro no lo es), puede consultar una tabla del artículo para determinar si una diferencia en el poder estadístico podría ser un problema. Esta diferencia de poder afecta sólo a una pequeña minoría de los casos.

En general, la elección entre los dos análisis es lazo. Si necesita comparar dos grupos de datos de Likert de cinco puntos, normalmente no importa qué análisis utilice.

Ambas pruebas proporcionan casi siempre la misma protección contra falsos negativos y siempre proporcionan la misma protección contra falsos positivos. Estos patrones son válidos para tamaños de muestra de 10, 30 y 200 por grupo.